Internet measurement: myths about Internet data

http://www.caida.org/outreach/presentations/

the work ‘problem’ implies an illusion:
that this problem I am having has definable limits.
everything runs into everything else
-- ‘i touch the earth and the earth touches me’

apr 2002
darpa nms N66001-01-1-8909

ucsd/sdsc/caida
kc@caida.org

what I mean by ‘myth’

- if you google for “Internet myths”, you’ll get lots of figments about Internet marketing/sociology, like
  - it’s cheap to do business on the web
  - advertising is flocking to the web in record numbers and will be its savior
  - you can give away the merchandise as long as you generate enough eyeballs because one day you will monetize those eyeballs
  - if you have a clever URL, they will come
  - people will never pay for content over the web
  - traditional advertising brings eyeballs which generates much traffic
  - people like to shop on the web (<-- that’s a good one)
  - it costs nothing to get a site up and running
  - the web is a reliable commercial activity
  - just you wait, profitability is right around the corner -- *http://www.thestreet.com/comment/wrongtactics/786636.html*

about these ‘myths’

these are not ‘myths’

since noone actually believes them

these are called fantasies

(people want them to be true

...or (more sustaining:) get return for convincing someone they’re true

myths: things people actually believe but that are wrong

fantasies vs myths

- **fantasies**
  - who believes:
    - marketing, advertising people, lawyers, consultant (consenting) adults
    - addictive drug users (in a low-ROE way)
  - who gets hurt:
    - marketing, advertising people? (no comment...)
- **myths**
  - who believes:
    - researchers, vendors, policymakers, journalists, secretary of defense potentially: marketing, advertising people, lawyers, consultant (consenting) adults
  - who gets hurt:
    - packets (dropped)
    - engineers (paged)
    - protocol developers (in worst case they invent stuff like atm, mpls)
    - grad students (useless dissertations, sub-employability, lost decades of youth)
    - economy (irrational speculation in capital markets -> global recession)

Internet myths relevant to engineering (about data)

- workload: (besides basic traffic growth fiction, which has been ludicrous)
  - level and nature of fragmented traffic
  - increase in flows as bandwidth grows
  - private addresses in core
  - mice vs. elephants
  - prevalence of encrypted passwords
  - applications can be identified (much less controlled)

Internet myths relevant to engineering (about data)

- performance:
  - – DoS attacks affect only large sites
  - – geography not correlated with latency
  - – DNS system performs well
  - – single router can’t trash the Internet
- topology:
  - – Internet topologies, object sizes follow power laws

Internet myths relevant to engineering (about data)

- routing:
  - routing tables reflect Internet topology
  - intra-country traffic stays there
  - AS path length is decreasing
  - small providers and multi-homing (more specifics) cause all the churn

why so many myths? no real measurement

Internet’s resistance to modeling and measurement

- evolution-based (good!) reasons
  - protocols, technologies, applications
    - independently developed and deployed
    - by no means synergistic
    - by all accounts rapid
    - ‘punctuated’ but no equilibrium
    - **“have done fine without modeling so far”**
    - ***(let’s wait till modeling is cheaper than bandwidth*****)**

...but simulation/analysis validation (& lately engineering/billing/security) needs data

- right granularities hard to come by
- measurement technology just not there
- argument for it also not there
- **“helps everyone”, but who pays?**
- losing battle?

measurement tools lack

- well-defined traffic metrics e.g. supporting SLAs or billing
- uniformly applied methodologies
  - varied topologies, equipment, ISP practices
- clear definition of measurement hypotheses or goals
- measurement scalability
- ability to explain phenomena
  - topology changes, routing loops, black holes
- relevance to actual ISP problems or mechanisms for repair
- communication of useful results

Internet’s resistance to measurement

- many would benefit
  - vendors, users, researchers, ISPs
- ISPs would bear cost
  - multiple media: atm, pos, dwdm, mpls
  - logistics/management
  - privacy implications
  - analysis/research obsolete after (before) done

...how to justify/accomplish measurement? (when market forces are torqued)

- alternatives:
  1) tools that positively affect an ISP's balance sheet
  2) regulatory intervention

what happened instead of measurement?

- from andrew odlyzko’s excellent “myth of Internet growth” study (nov 2000) plus great assessment (...) of larry roberts caspian.goo last month
  - ‘traffic doubling every 90 days’
    - maybe for a few months in 1995-1996
    - in reality, no real data since 1995 (nsfnet sunset)
    - more like every 12-18 months for rest of 1990s
  - financial markets (at least in US) believed (bubbly!) estimates

what happened instead of measurement?

- over 6 years, that means a factor of 16 million
  - assume (generously) 500M users, 1.5Mbps per user **around the clock**
  - **and yet we're mostly still using 28k modems, & only for an hr/day, & avg 5k bits/sec even then**
  - **the math just does not work out**
- it took 5 years for true traffic growth data to finally manifest itself (since providers would not release data, if they even had it)
  - via other metrics (hardware and bandwidth sales) required in annual reports to SEC (closest we have to an Internet Measurement Commission)
- **that's actually an embarrassingly pathetic willingness to ignore real data (or just invent it)**

living in a mythical world: tradeoffs

- costs
  - tech stock bubble? (hey infinite demand is infinite jnpr stock price)
    - really takes new technologies a decade to penetrate
    - web was exception (when it was young/free), Internet is not
  - retarded technical developments
  - negligence of what users want and are likely to get
    - community gets mired in sub-necessary QOS hubbub, ATM, GMPLS

living in a mythical world: tradeoffs

- benefits
  - – unparalleled platform for innovation
  - – open standards, rapid development of new services
  - – big empty pipes were key factor in supporting [r]evolution
    - pipes wouldn’t be empty for grad students (napster, kazaa) if the myths had been true

living in a mythical world: tradeoffs

- lessons
  - 25 year contracts for pipes should be amortized over 3 years
  - come to terms with a much looser definition of ‘capacity planning’
  - simplify engineering (atm/sonet --> IP over WDM, GigE)
  - **(first commandment: Thou Shalt Get Rid of Layer Goo)**

living on borrowed time in a mythical world

(opportunity costs of measurement)

- three ‘waves’ of Internet applications / usage
  - ***first wave:*** shared (remote) use of computers
    - telnet, email, ftp
  - ***second wave:*** client/server model, formatted languages
    - web
  - ***third wave:*** collaborative, peer-to-peer, interactive
    - napster, imesh, kazaa, gaming, video

living on borrowed time in a mythical world

- emergence of third wave (‘ngi’) will require more real-time interaction with and reaction to network status
- the growth of these applications will be self-limiting (by user frustration with performance) unless we have either:
  - a better grip on measurement
    - either done by the applications themselves (e.g., vat)
    - or via some other middleware aspect of the infrastructure
  - or no service-affecting queueing anywhere in the network
    - seems unlikely, even with lots of empty pipes

four areas of measurement (and thus myths)

- workload characterization (passive)
- topology (mapping, path dynamics)
- performance evaluation (active, passive)
- routing (dynamics)

caida focuses on

- measurement tools (prototypes)
- macroscopic (or macroscopically relevant) analyses
- identifying priorities and obstacles

workload measurement: dag oc48 capture card

- current oc48mon system (prototype at MFN in SJC, subc/collab. w U. Waikato
  - captures 1M packets/sec to disk (40% util. link)
  - provides highly accurate timestamping
  - .5Mp, 1Gbps (125MB/sec) each direction
  - avg pkt size 370, 590 bytes (210k, 240k ptks/sec)
  - 64 bytes/record -> 6-9x compression over link load
  - problems: bursts of small packets cause machine thrash
  - http://dag.cs.waikato.ac.nz/

workload measurement: dag oc48mon card

- upgrading oc48mon this qtr to house (bigger) Dag4.10 cards
  - dual-Pentium (Intel) processor on tyan S2510
  - 1Gb of RAM
  - floppy, cdrom
  - IDE/ATA disk drive (40Gb min)
  - 6 SCSI Ultra/160 disks, 3/each SCSI channel each 18Gb min
  - 4U rack mountable chassis

this will get us One Hour (and just barely, and ~50Gb) (MFN SJC 76 min 020:00 PDT 5 aug 2001 ==> 32Gb)

workload measurement: dag+coral oc48mon

- **unique**
  - first and only OC48 flow monitor worldwide
  - caida’s public tools analyze data without modification
- **software implemented**
  - CoralReef, NeTraMet, custom routine (CAIDA)
  - other custom/enhanced routines by U. of Waikato, others
  - darpa/nsf/caida members funded
- **software, data analysis, viz tools all prototypes**
  - commercial spinoff for the cards (www.endace.co.nz)
- **but btw backbone core now needs oc192/oc768 monitoring**
  - currently no such project exists (someone tell homeland security office)

workload myth: mice vs elephants

- myth: 10% of flows contribute 90% of total traffic on a link
- **data:**
  - sometimes true for bytes
    - if the link has KaZaa-type stuff
  - never true for packets
    - in any traces we've studied
  - actual proportion of traffic (bytes or packets) covered by 90% of streams can change rapidly following changes in the applications/protocols mix

--> need to measure proportions before making assumptions

--> need longer traces!

measurements: analysis

- use CoralReef software suite
  - http://www.caida.org/tools/measurement/coralreef/
- obtain quantitative parameters of captured traffic:
  - Byte rates and Packet rates
  - Flows
    - Flow = (src IP, src port, dest IP, dest port, protocol)
- use NetGeo tool to map src/dst IP addresses to ASes and countries
  - http://www.caida.org/tools/utilities/netgeo/
- consider various aggregations of traffic:
  - applications
  - ASes
  - countries

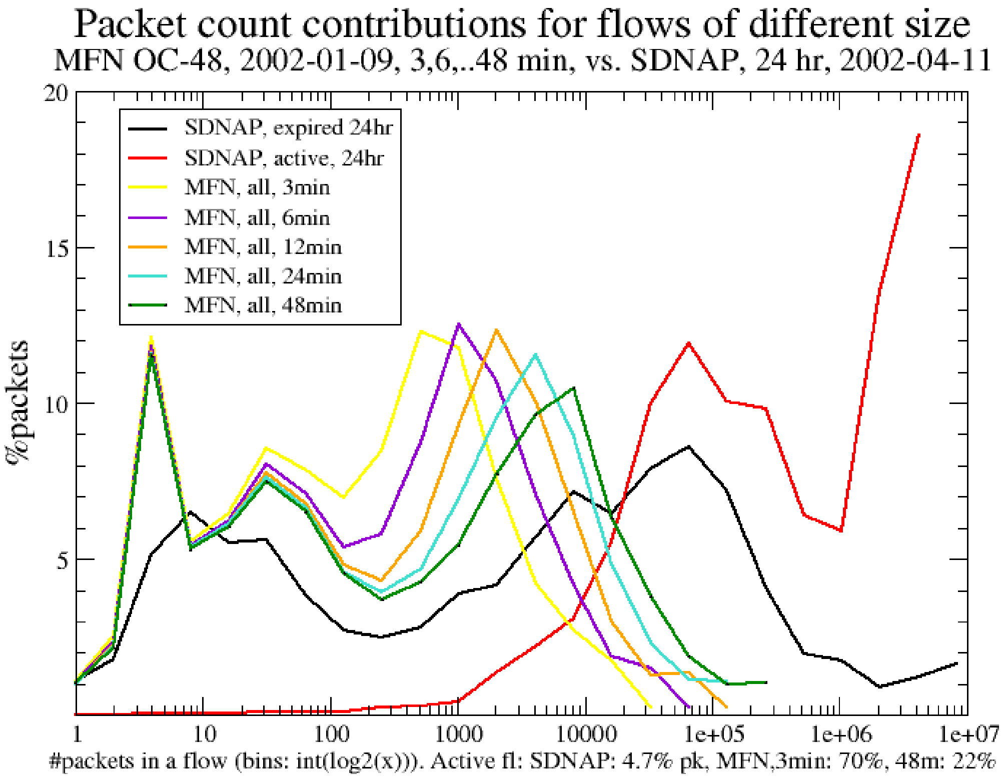
workload: mice vs elephants

- **two modes of Internet usage (interactive, downloads)**
  - boundary between modes is ~300 packets (0.5 Mbytes)
- **most flows on the left (by far), most packets on the right (by far)**
- **for a 24 hour (sd) trace, 4.7% packets are in still-active flows**
  - 50% packets are in flows with >8192 ppkts; max. flow: 9Mpkts max. active flow: 5Mpkts
- **for a 3 min (sjc) trace, 70% pkts in still-active flows**
- **for each 2% in sample duration, 2% in max of pkt/flow**
- **convergence nowhere in sight**

➔do not study flow sizes with less than 24 hrs of data

workload: mice vs elephants

➔ do not study flow sizes with less than 24 hours of data

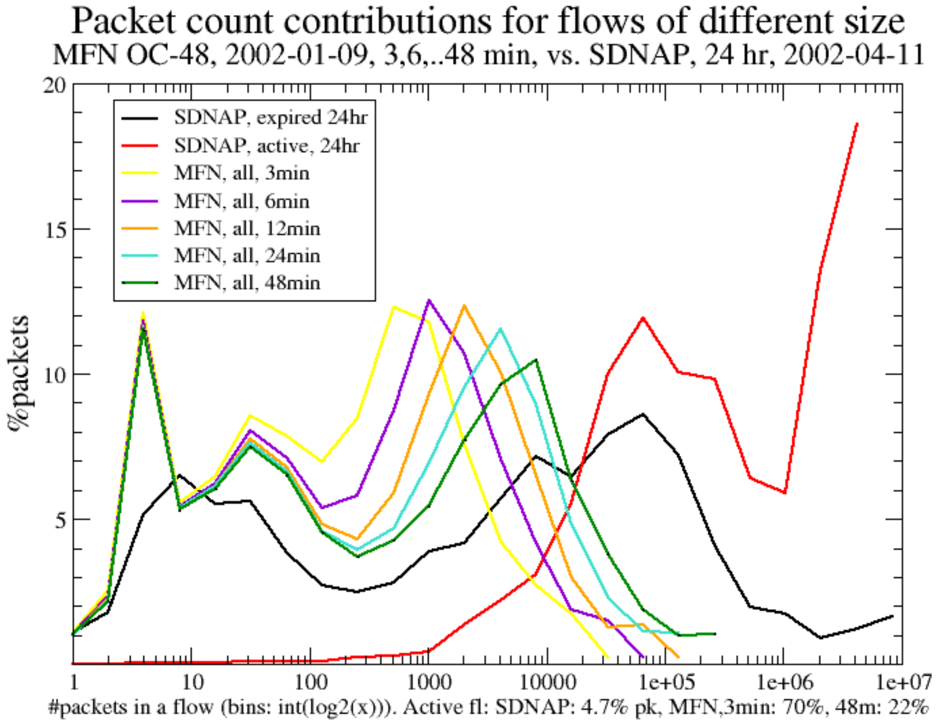


workload: mice vs elephants

(generally, we do not yet know what we’re talking about)

→ **but we know not to study flow sizes with less than 24 hours of data**

→ **btw, nobody has 24 hours worth of useful data (we’re $5M away)**



workload myths: prevalence of IP fragmentation

- myth: there is no fragmented traffic
- data: while true that only a small percentage (0.09% - 1.6%) of traffic is fragmented, this number is growing. Some protocols, for example IGMP, have fragmented traffic far exceeding non-fragmented traffic.

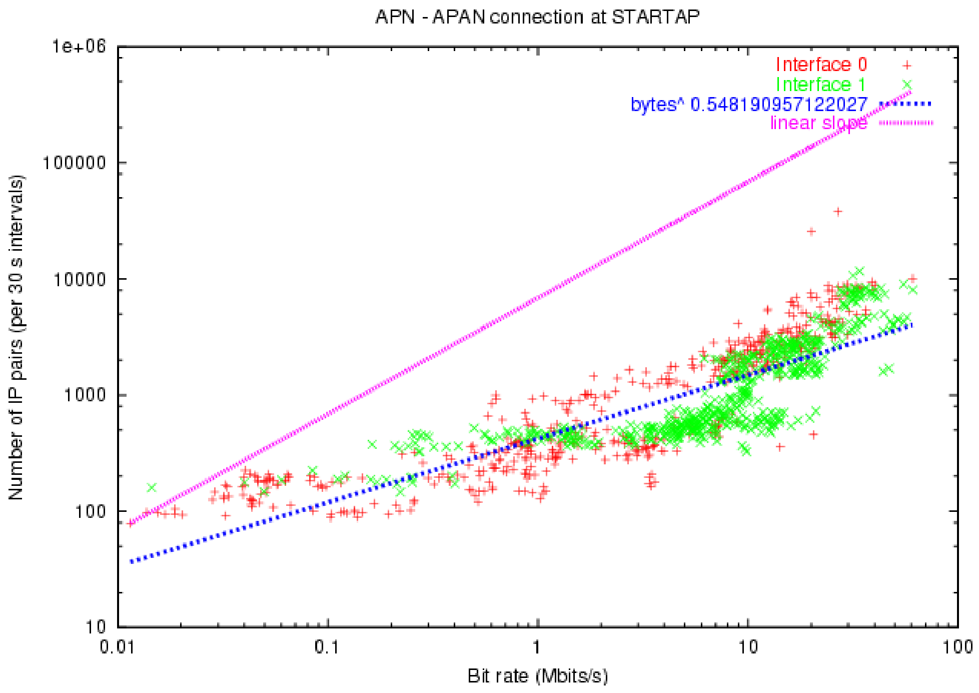
workload myths: prevalence of IP fragmentation

- myth: fragmented traffic exists only on LANs
- data: we've monitored it on aggregated exchange points and backbone links.
- myth: tcp traffic is never fragmented
- data: while tcp traffic is fragmented much less frequently than other protocols due to path MTU discovery, we monitored 0.009% by packets (0.019% by bytes) of fragmented tcp traffic and a majority of fragmented tunneled traffic is TCP!

workload myths: prevalence of IP fragmentation

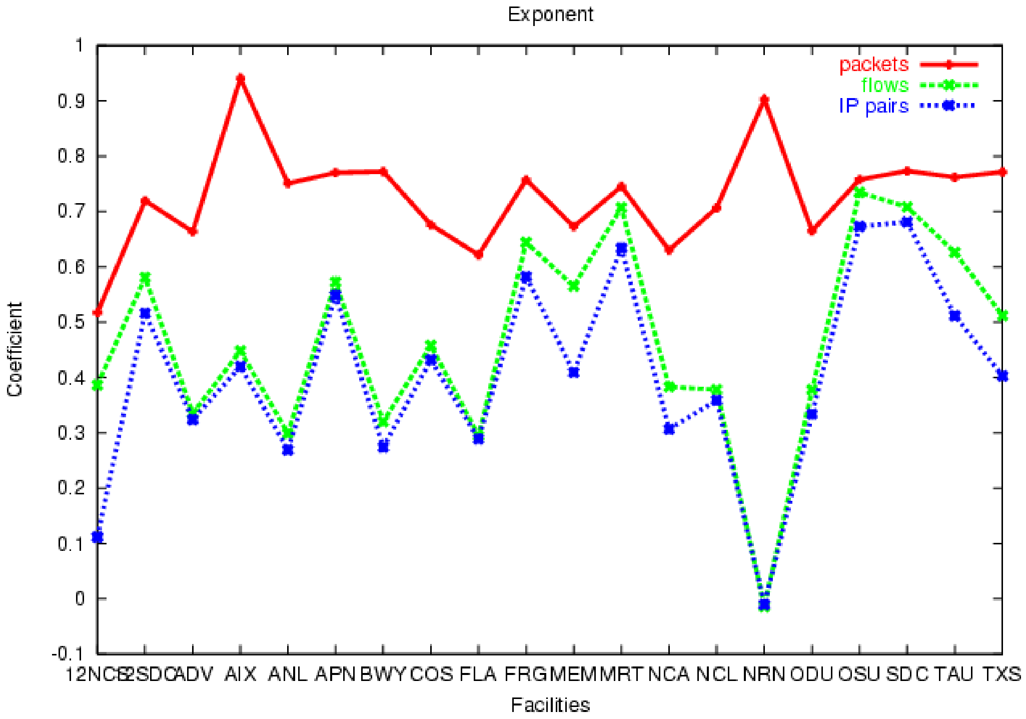
- myth: NFS causes all (or almost all) fragmented traffic
- data: tunneled traffic (IPENCAP, IPIP, GRE, UDP L2TP), ICMP, and RealMedia all caused more fragmented traffic than NFS (0.1%)

workload myth: # host pairs increases as square of bandwidth



- data: growth much slower than linear
  - (20 academic sites over 4 years, 2900 nlanr/moat traces) growth spans 4 orders of magnitude

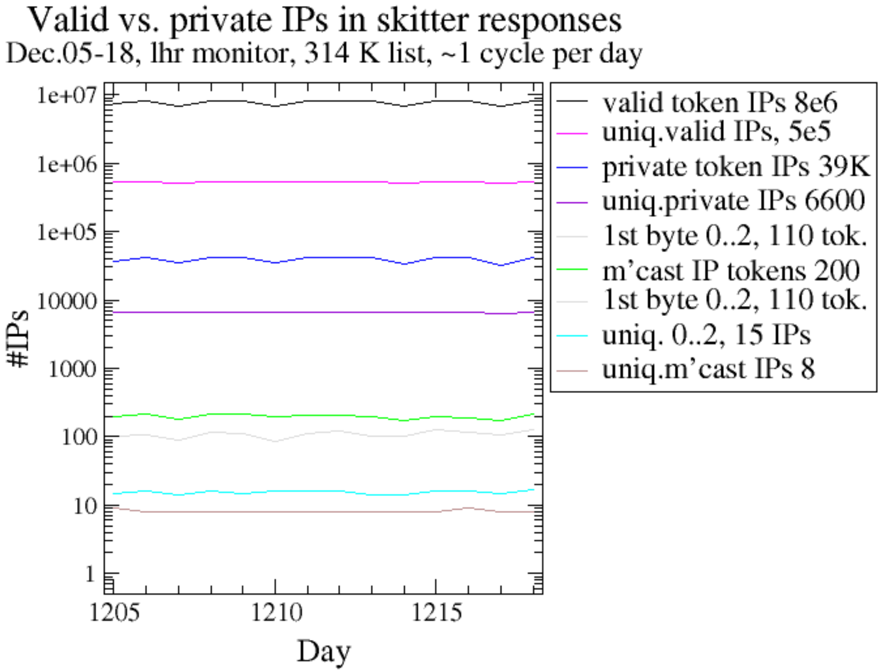
workload myth: host pairs increase as square of bandwidth (2)



- **data:** for all monitored facilities:
  - ***pkts vs. bit rate*** - growth is nearly linear (power $^{a\sim1}$)
  - ***flows and IP pairs vs. bit rate*** - grow as square root (a ~ 0.5)

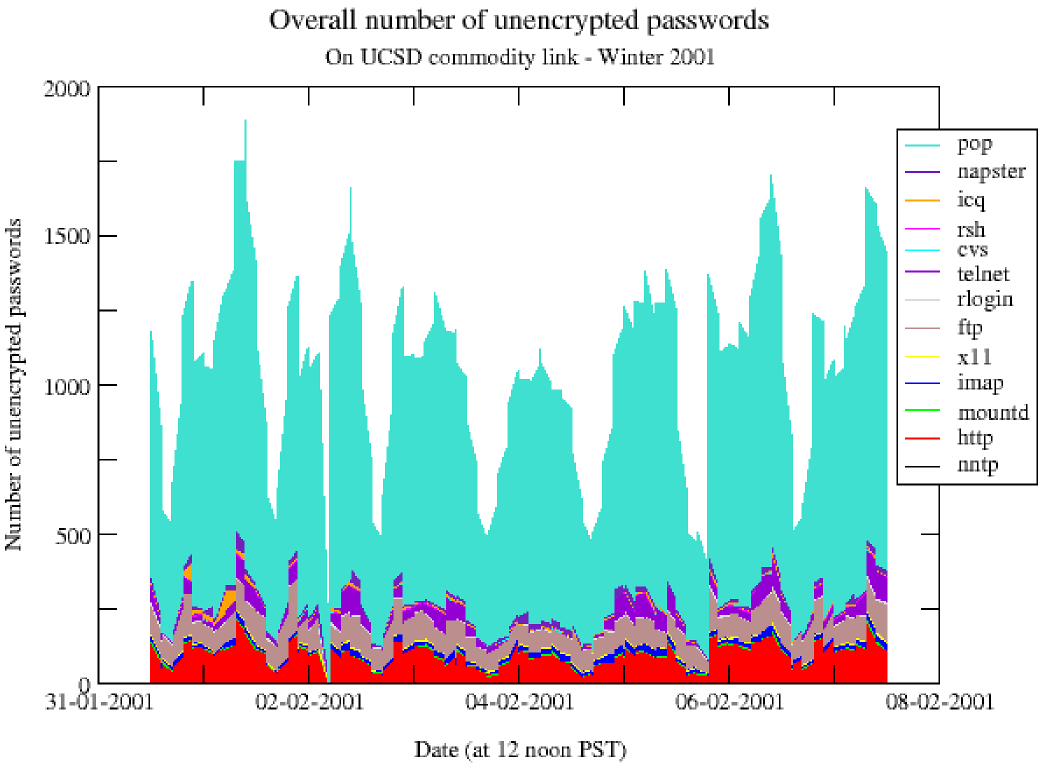
workload myth: private addresses do not appear in the core

- **data:** private addresses appear all over the place
  - including (consistently) in queries to root name servers
  - as do multicast and other ‘shouldn’t be seen’ junk
  - Broido’s 1st Law: ‘what should not be seen in the Internet will appear 1% of the time’



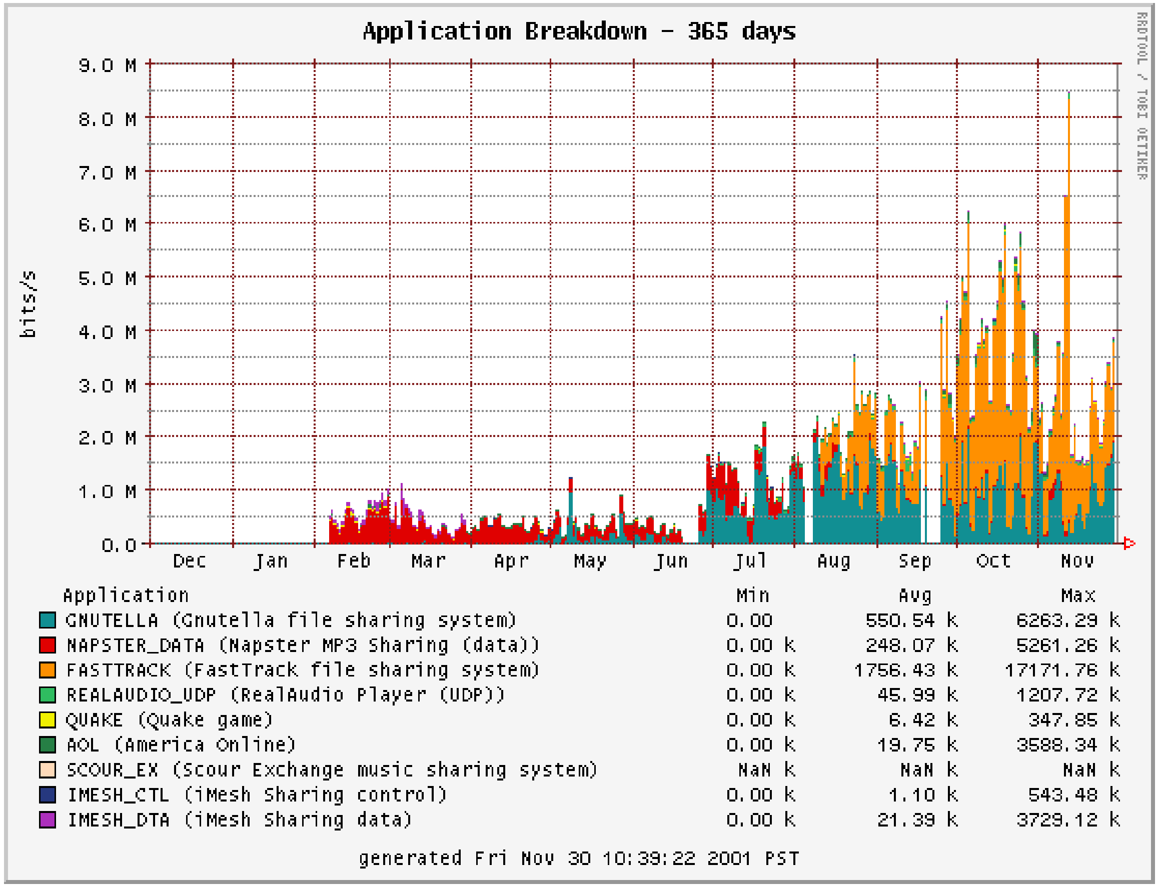
workload: prevalence of encrypted passwords

- myth: unencrypted passwords mostly gone
- data: most unencrypted passwords are from one source: POP
  - why aren't folks using APOP? (authentication already provided)
  - mere existence of an encryption technology is no guarantee of its adoption



workload myth: US govt can stop file sharing

- admit it’s in fantasy category (myth might also be stated as ‘currently there is no killer app’)



Application	Min	Avg	Max
🟦 GNUTELLA (Gnutella file sharing system)	0.00	550.54 k	6263.29 k
🟥 NAPSTER_DATA (Napster MP3 Sharing (data))	0.00 k	248.07 k	5261.26 k
🟧 FASTTRACK (FastTrack file sharing system)	0.00 k	1756.43 k	17171.76 k
🟩 REALAUDIO_UDP (RealAudio Player (UDP))	0.00 k	45.99 k	1207.72 k
🟨 QUAKE (Quake game)	0.00 k	6.42 k	347.85 k
🟩 AOL (America Online)	0.00 k	19.75 k	3588.34 k
⬜ SCOUR_EX (Scour Exchange music sharing system)	NaN k	NaN k	NaN k
🟦 IMESH_CTL (iMesh Sharing control)	0.00 k	1.10 k	543.48 k
🟪 IMESH_DTA (iMesh Sharing data)	0.00 k	21.39 k	3729.12 k

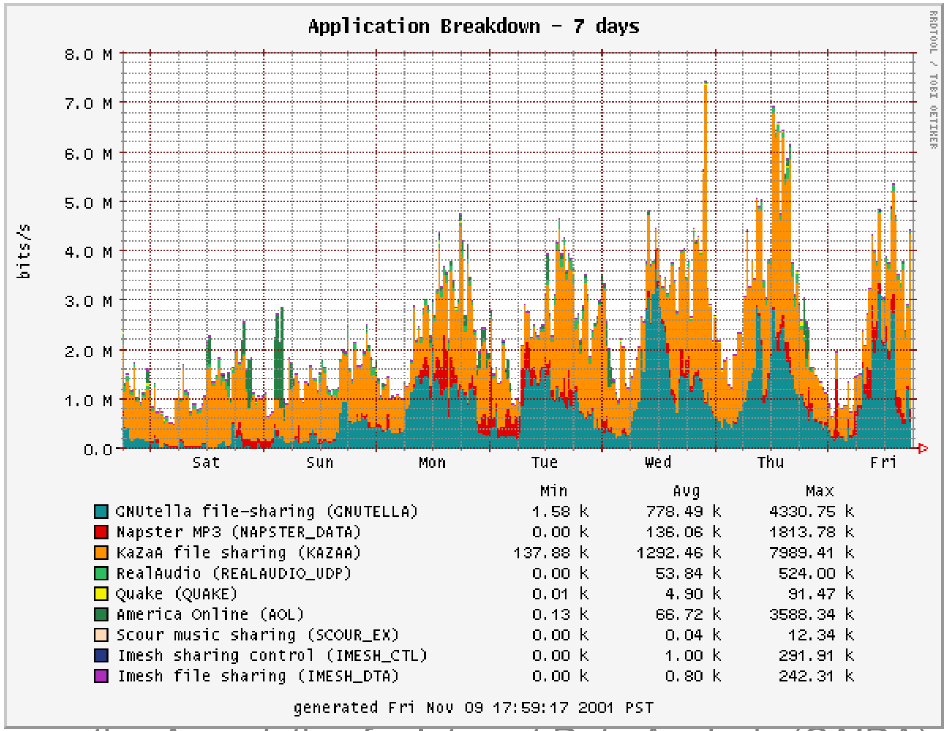
generated Fri Nov 30 10:39:22 2001 PST

in an expanding system, such as a growing organism, freedom to change the pattern of performance is one of the intrinsic properties of the organism itself

workload myth: govt can stop file sharing / no killer app (2)

"how do you know when something is a 'killer app'? when every university tries to stop it and can't. that's how you know it's a killer app. that it takes a federal judge to threaten to put you in jail if you don't stop. THAT's how you know it's a killer app!"

- eric schmidt, keynote for dns navigation workshop



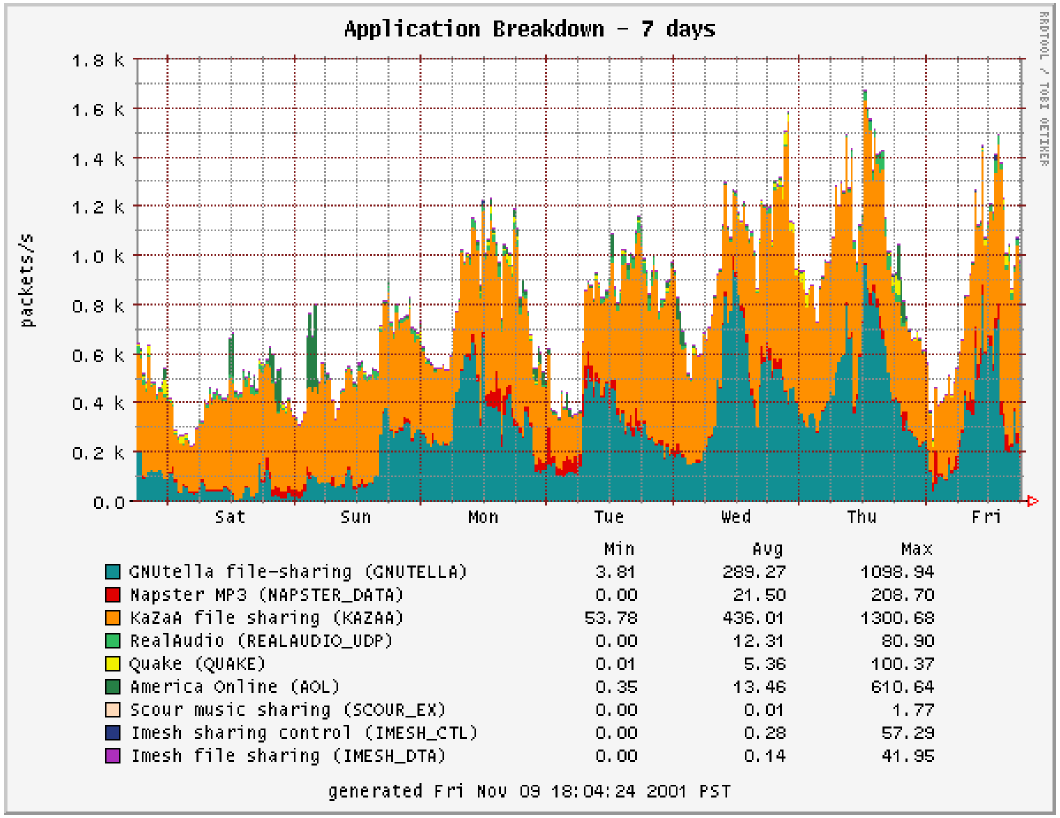
	Min	Avg	Max
🟦 GNUtella file-sharing (GNUTELLA)	1.58 k	778.49 k	4330.75 k
🟥 Napster MP3 (NAPSTER_DATA)	0.00 k	136.06 k	1813.78 k
🟧 KaZaA file sharing (KAZAA)	137.88 k	1292.46 k	7989.41 k
🟩 RealAudio (REALAUDIO_UDP)	0.00 k	53.84 k	524.00 k
🟨 Quake (QUAKE)	0.01 k	4.90 k	91.47 k
🟩 America Online (AOL)	0.13 k	66.72 k	3588.34 k
⬜ Scour music sharing (SCOUR_EX)	0.00 k	0.04 k	12.34 k
🟦 Imesh sharing control (IMESH_CTL)	0.00 k	1.00 k	291.91 k
🟪 Imesh file sharing (IMESH_DTA)	0.00 k	0.80 k	242.31 k

generated Fri Nov 09 17:59:17 2001 PST

Cooperative Association for Internet Data Analysis (CAIDA)

workload myth: govt can stop file sharing / no killer app (3)

- in case you thought it was just huge packets sneaking in
- also note similarity to gopher/web transition (patent/port# control)
  - (not that anyone would know via measurement... ask Internet historian)



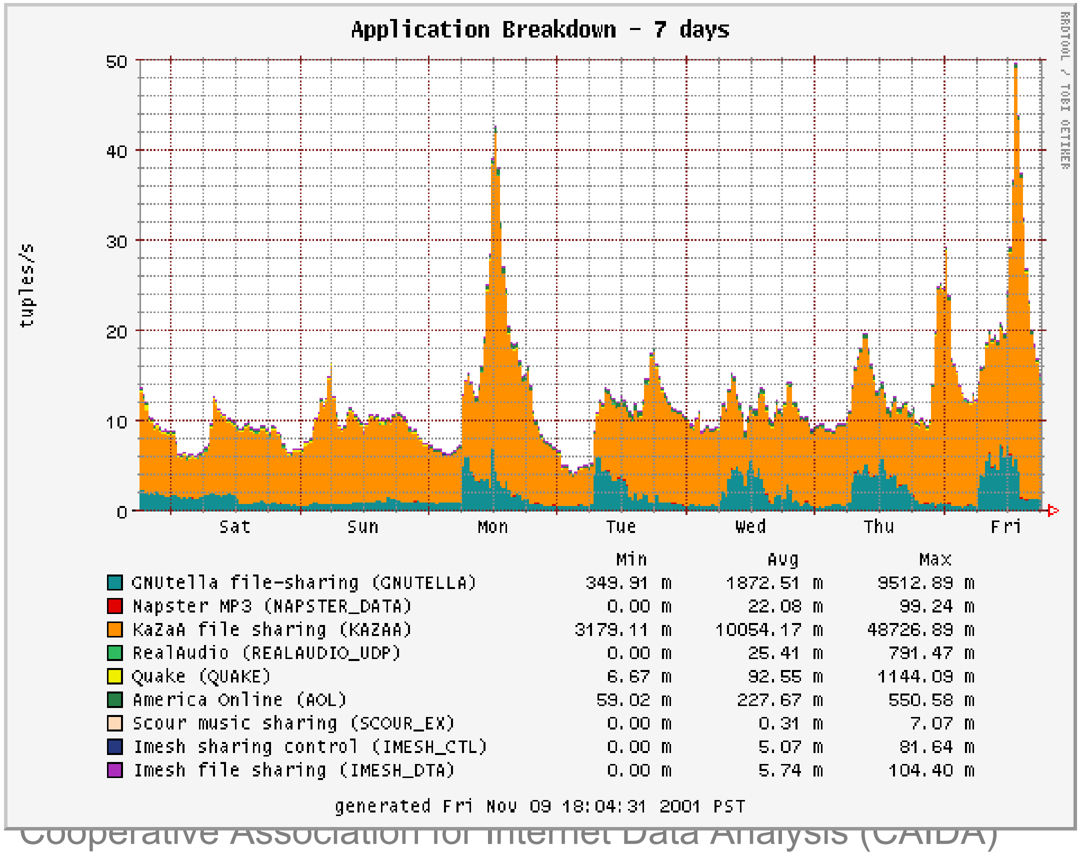
	Min	Avg	Max
🟦 GNUtella file-sharing (GNUTELLA)	3.81	289.27	1098.94
🟥 Napster MP3 (NAPSTER_DATA)	0.00	21.50	208.70
🟧 KazaA file sharing (KAZAA)	53.78	436.01	1300.68
🟩 RealAudio (REALAUDIO_UDP)	0.00	12.31	80.90
🟨 Quake (QUAKE)	0.01	5.36	100.37
🟩 America Online (AOL)	0.35	13.46	610.64
⬜ Scour music sharing (SCOUR_EX)	0.00	0.01	1.77
🟦 Imesh sharing control (IMESH_CTL)	0.00	0.28	57.29
🟪 Imesh file sharing (IMESH_DTA)	0.00	0.14	41.95

generated Fri Nov 09 18:04:24 2001 PST

Cooperative Association for Internet Data Analysis (CAIDA)

workload myth: govt can stop file sharing / no killer app (4)

- in case you thought it was just a few punks
  - compare how different apps affect network... especially bytes vs. tuples
  - gnutella/fasttrack: both big flows; fasttrack (kazaa): lot more connections



	Min	Avg	Max
🟦 GNUtella file-sharing (GNUTELLA)	349.91 m	1872.51 m	9512.89 m
🟥 Napster MP3 (NAPSTER_DATA)	0.00 m	22.08 m	99.24 m
🟧 KaZaA file sharing (KAZAA)	3179.11 m	10054.17 m	48726.89 m
🟩 RealAudio (REALAUDIO_UDP)	0.00 m	25.41 m	791.47 m
🟨 Quake (QUAKE)	6.67 m	92.55 m	1144.09 m
🟩 America Online (AOL)	59.02 m	227.67 m	550.58 m
⬜ Scour music sharing (SCOUR_EX)	0.00 m	0.31 m	7.07 m
🟦 Imesh sharing control (IMESH_CTL)	0.00 m	5.07 m	81.64 m
🟪 Imesh file sharing (IMESH_DTA)	0.00 m	5.74 m	104.40 m

generated Fri Nov 09 18:04:31 2001 PST

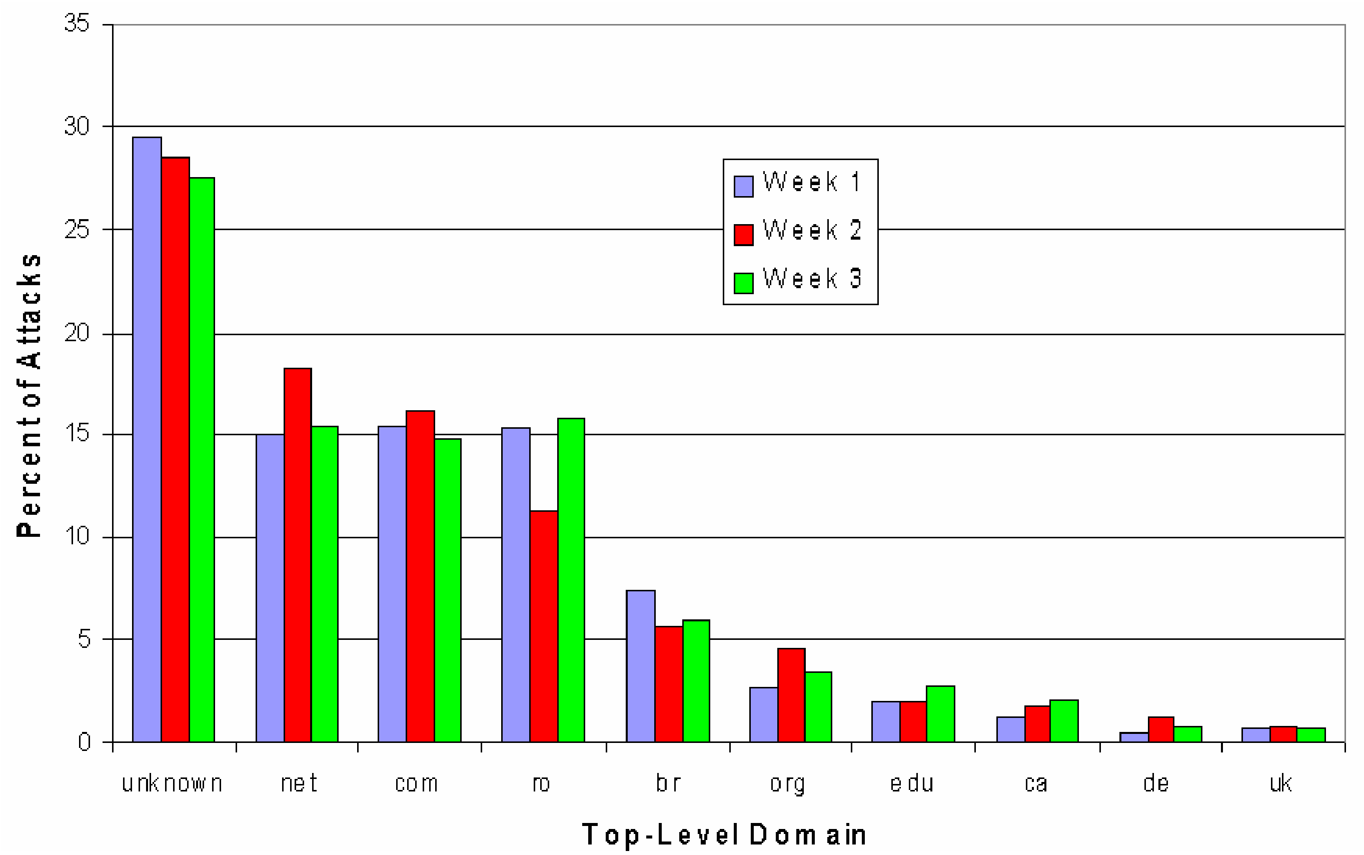
performance myth: DoS attacks

- myth: flooding DoS attacks only affect large commercial sites, are long in duration and at extremely high rates
- **data:** >12,000 attacks against >5,000 targets in 3 weeks
- **~20-60** attacks occurring at all times
- **80%** of attacks last ***less than an hour***, a few lasted ***3 weeks***
- **70%** of attacks ***<1,000pps***, some over ***600,000pps***
- **10-20%** of attacks to home machines (***cable, dsl,dialup***)
- **5%** of attacks target infrastructure (***routers, dns servers)***

(usenix 2001, david,colleen@caida.org, stefan,geoff@ucsd.edu)

performance myth: DoS attacks (2)

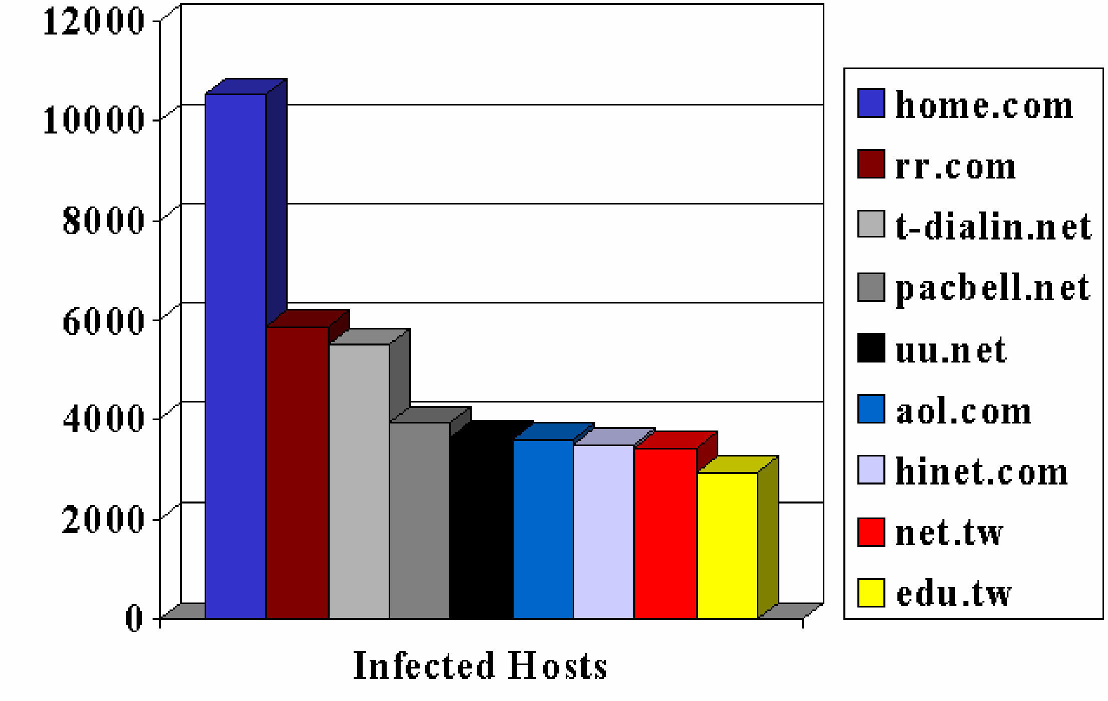
- romania and brazil have disproportionate number of infected hosts
- other domains have roughly same ratio of infected/total machines



Cooperative Association for Internet Data Analysis (CAIDA)

performance myth: worm spread

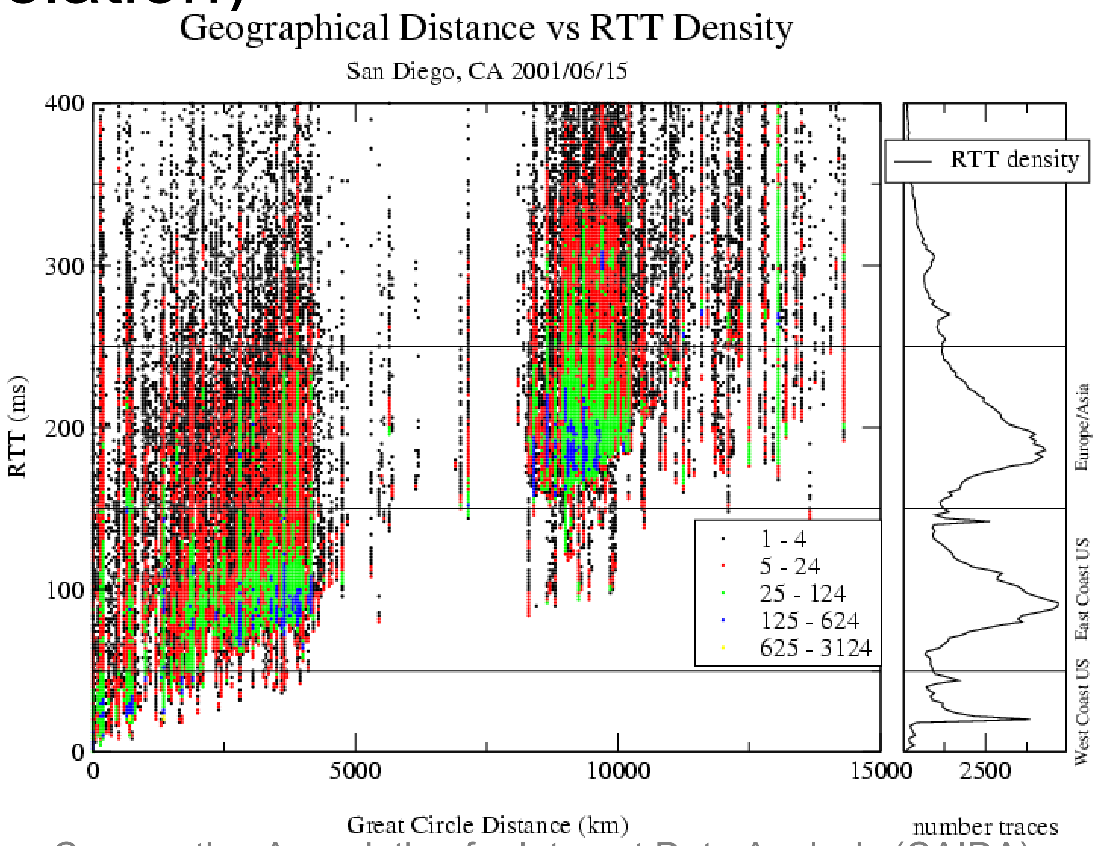
- 40% of all hosts infected (first round CodeRed) lacked reverse DNS records, so we were unable to determine their hostnames
- ISPs providing connectivity to home and small-business users had the most infected hosts
- machines maintained by home/small-business users (i.e. less likely to be maintained by a professional sysadmin) are an important aspect of global Internet health



Cooperative Association for Internet Data Analysis (CAIDA)

performance myth: geography not correlated w/latency

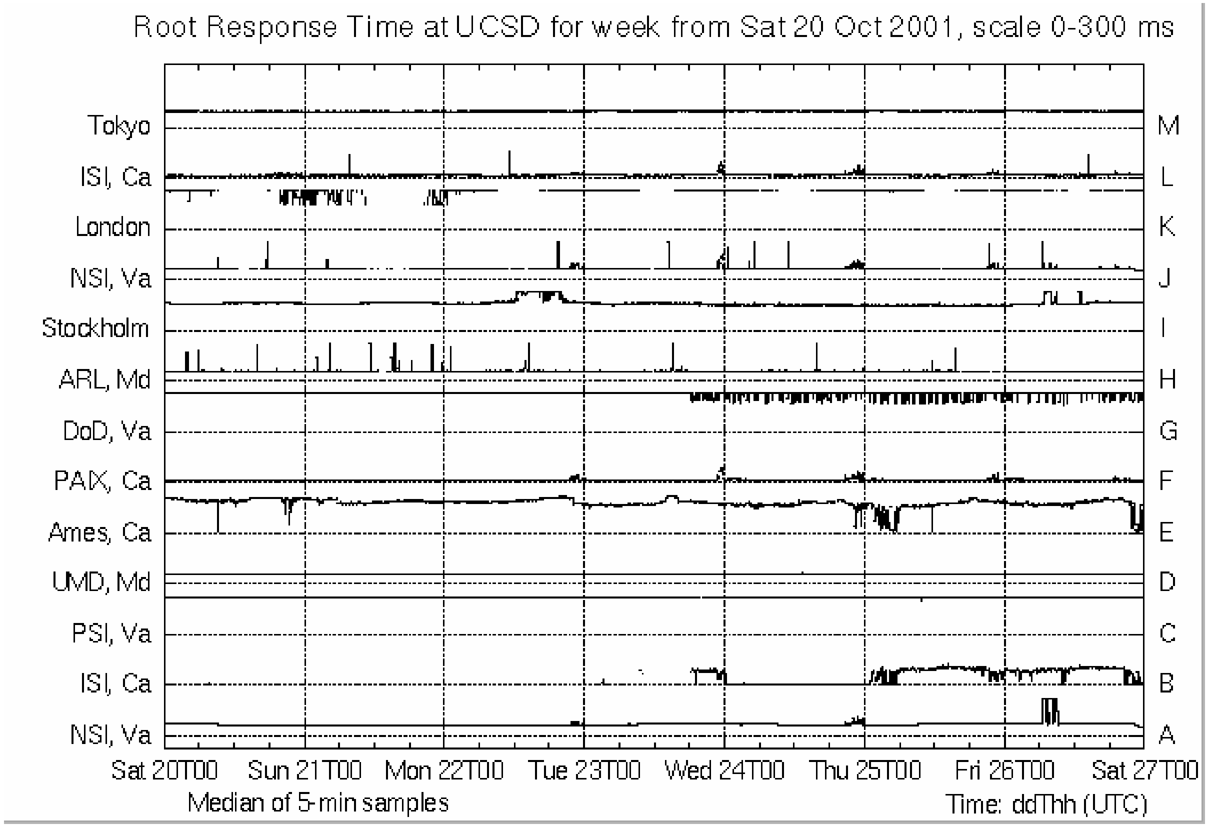
- data: rtt densities from san diego (strong correlation)



Cooperative Association for Internet Data Analysis (CAIDA)

performance myth: root DNS system performs well

- data: 8 of the 13 root servers perform well, so users don't notice the poor performance of the other five (actually gTLDs do better)



Cooperative Association for Internet Data Analysis (CAIDA)

performance myth: the DNS system performs well

- error taxonomy: bogus A queries to root name servers for a few hours at f-root in 2001
  - A queries ask for the IP address of a hostname
    - not supposed to be ‘in theory’
  - malformed A queries were 14% of the load at f-root
    - guilty: microsoft: Win2k resolver, viruses (win95/98/nt), macOSX resolver
    - asking for the IP address of an IP address
  - 20% of queries asking for non-existent TLD
    - lots of internal Microsoft names (active directory)
    - lots ending in .local, .localhost, .workgroup, .msft, .domain, etc
  - hard to track down, nameservers just relay clients queries
  - can’t see back to the actual client that asked the question

performance myth: single router can’t trash the Internet

(‘certainly not by accident’)

(hint: just need to trash 13 hosts to effectively trash the Internet)

just one example: microsoft’s feb 2001 dns woes

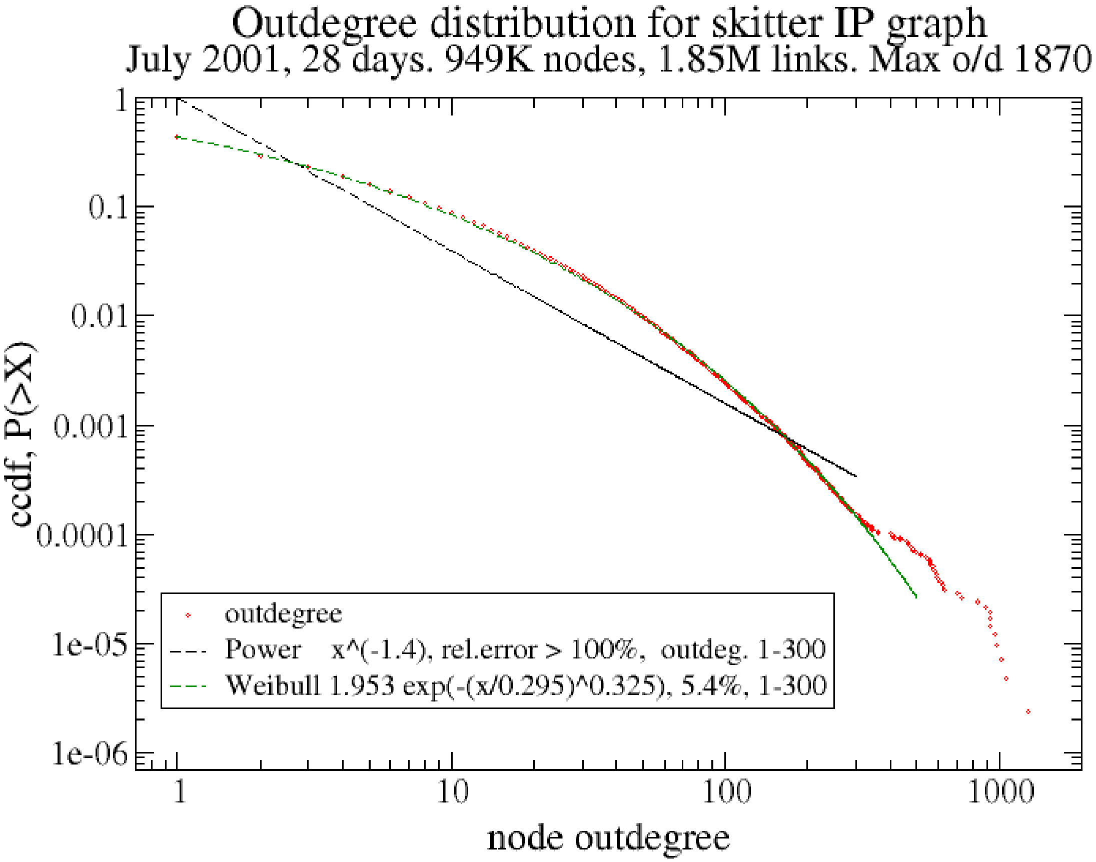
- microsoft’s 4 authoritative nameservers visible to world on one subnet (and now all you need is a comma in the wrong place)
- misconfigured router upstream of that subnet
- TTL for their names set to 2 hours
- started timing out of people’s caches
- query load at the roots started climbing
- microsoft nameservers don’t do negative caching

performance myth: single router can’t trash the Internet (continued)

- microsoft properties are usually about 6k queries/hour (0%)
  - increased to 25% of the load at f-root
- data: prominent site w/DNS problems affects whole Internet
  - cf. 9/11 cnn.com queries to roots were sustainable because of caching
  - this only a tiny piece of the root-server workload damage found

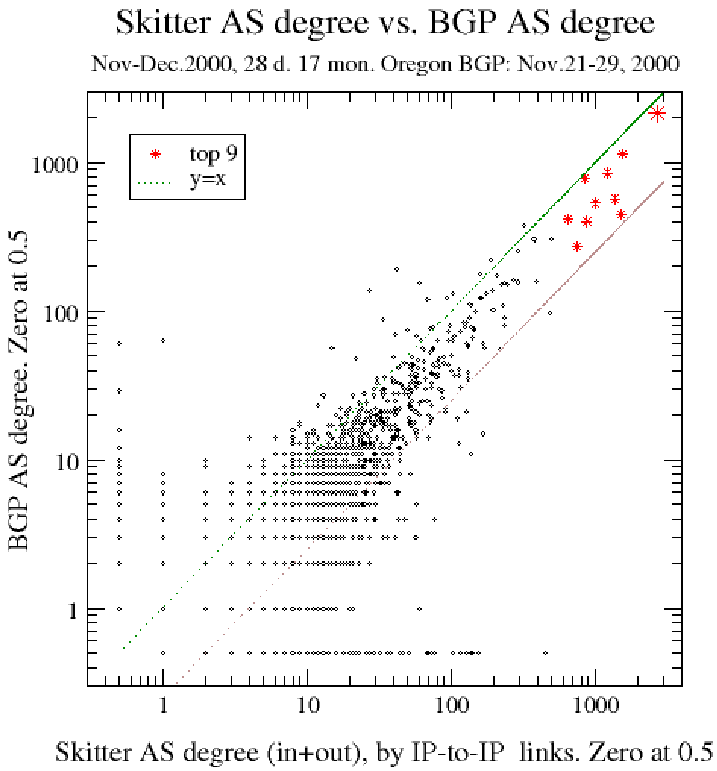
topology myth: outdegree distrib. follows power law

- **data:** distribution follows Weibull far better than power law



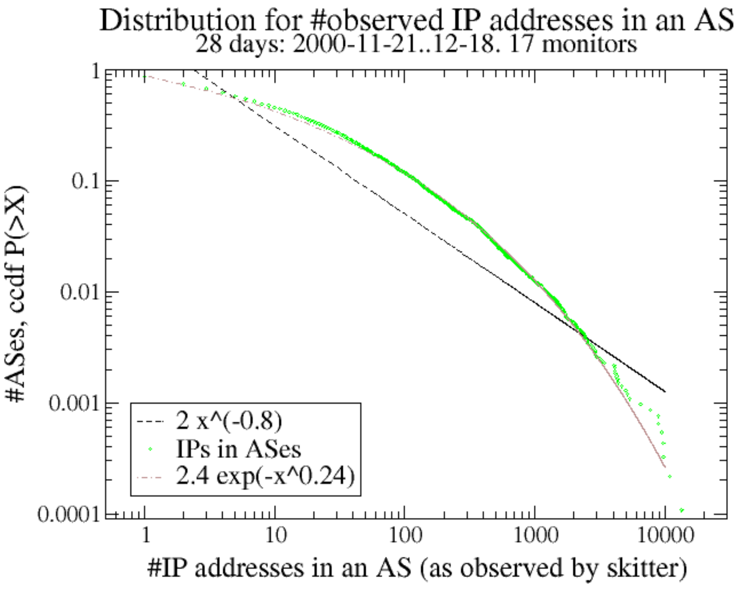
Cooperative Association for Internet Data Analysis (CAIDA)

topology myth: routing table data reflects topology



- **data:** **even the best available inter-domain routing (BGP) data serves as weak substitute for IP probed topology data (and yet this BGP data is an essential tool for sensible macroscopic Internet topology analysis)**

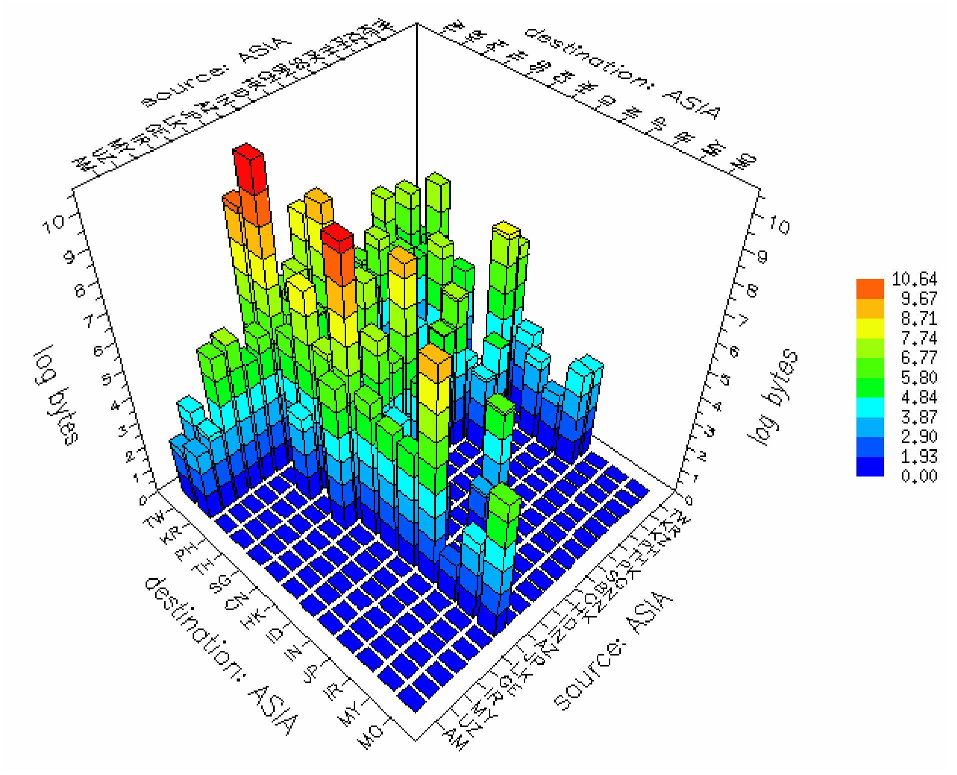
topology myth: Internet object sizes follow power law



- **data:** Internet graphs are closer to Weibull than to power functions
- $P(X>x) = a^{\mathbf{(-(x/b^\wedge c)}}$
- decreases faster than power function, slower than exponential

routing myth: intra-country traffic stays there

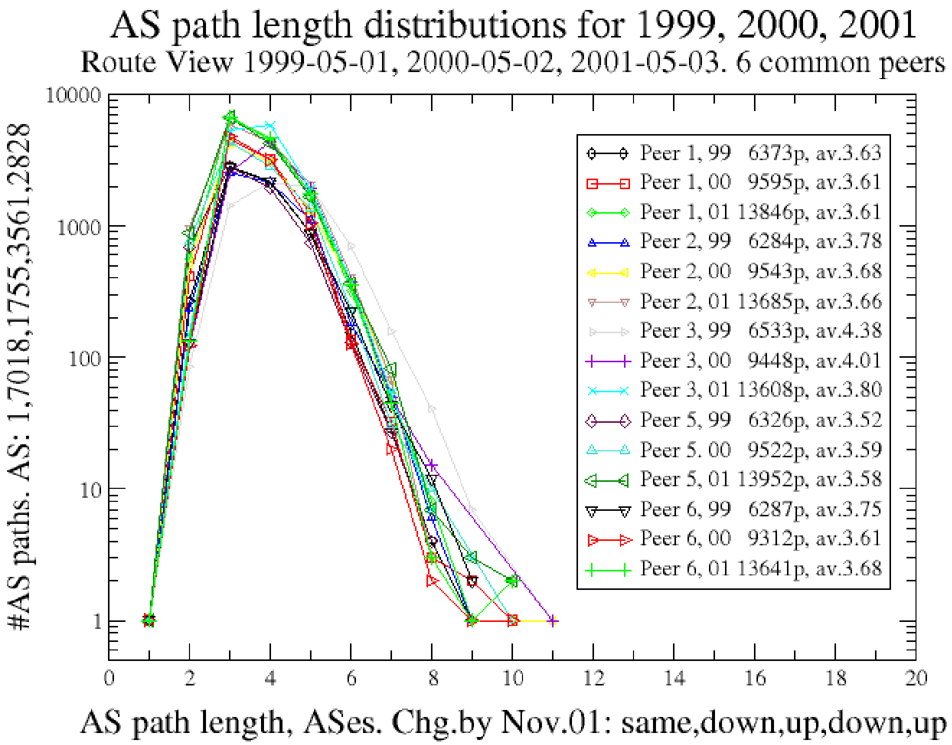
- **data: significant asia←→asia traffic goes thru san jose**
  - includes even some country traffic (e.g. .jp->.jp, .tw->.tw)



Cooperative Association for Internet Data Analysis (CAIDA)

routing myth: AS path length is decreasing

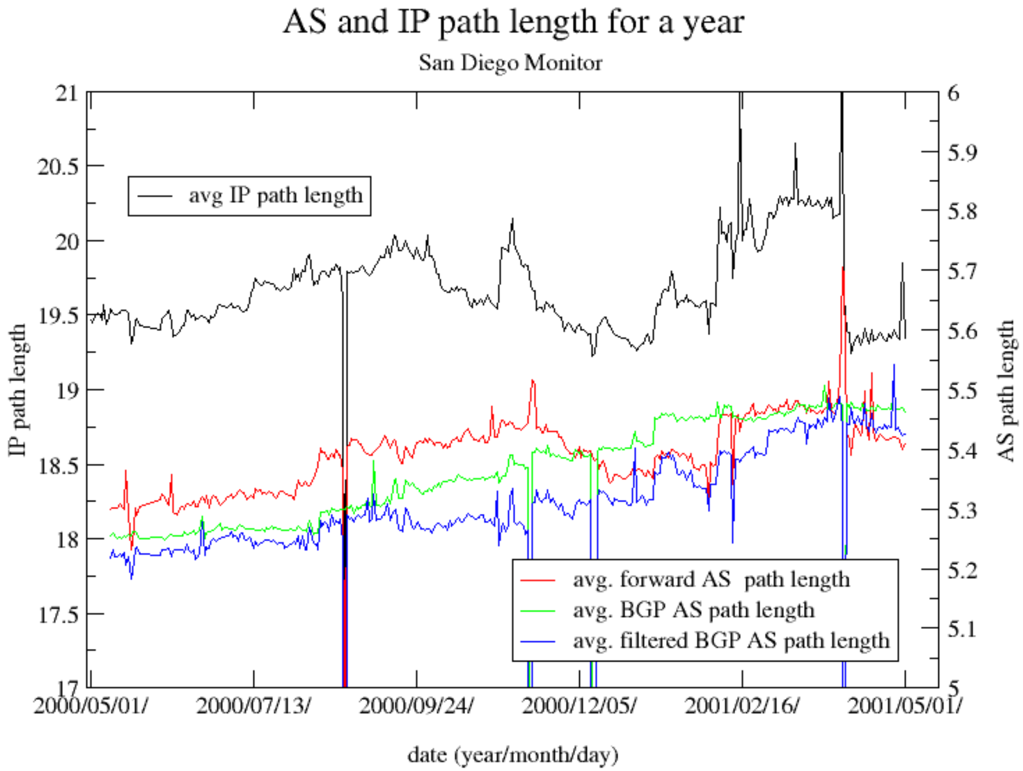
- **data: since 1999, many AS paths have changed either way**
  - average length decreased and increased for many ASes
  - change in the average AS path length is insignificant



Cooperative Association for Internet Data Analysis (CAIDA)

routing myth: AS path length is decreasing (cont)

- **data:** if anything, it’s increasing



Cooperative Association for Internet Data Analysis (CAIDA)

routing myths: causes of growth & instability of routing system

- myth: route table growth exponential
- **data:** global prefixes grew 4% may->nov 01; 37% in nov00-01 (RouteViews)
- myth: peering richness is growing (see last slide)
- **data:** link/node ratio (avg degree), peering richness, and churn did not significantly change in 2000-2001, although lots of changes within ASes

routing myths: causes of growth & instability of routing system (2)

- **myth:** small ISPs & multihoming cause growth and/or churn
- **data:** number of non-transit multihomed ASes grew from 35% to 37% in 2000-2001, but their share of global routes remained stable at around 30%
- **data:** new address announcements & deaggregation of existing prefixes were major sources of new prefixes between nov00-may01
- **data:** most routing instability (w/drawal/reannounce events) in late 2001 contributed by a few .gov networks, developing country telecoms, & major backbone ISPs, although backbone providers routes are relatively stable on per-prefix basis.
- **data:** instability caused in part by deaggregated routes leaking out originating AS, and by relatively short-lived transient announcements. ('small multihomers' contribute negligibly, at least on bi-hourly scale)

Internet myths relevant to engineering (about data)

- (besides basic traffic growth fiction)
  - level and nature of fragmented traffic
  - increase in flows as bandwidth grows
  - private addresses in core
  - mice vs elephants
  - prevalence of encrypted passwords
  - applications can be identified (much less controlled)

Internet myths relevant to engineering (about data) (continued)

- **performance myths:**
  - DoS attacks affect only large sites
  - geography not correlated with latency
  - DNS system performas well
  - single router can’t trash the Internet
- **topology myths:**
  - Internet topologies, objects sizes follow power laws
- **routing:**
  - routing tables reflect Internet topology
  - intra-country traffic stays there
  - AS path length is decreasing
  - small providers and multi-homing (more specifics) cause all the churn

why so many? no real data/measurement

conclusions

- we shed doubt on (too many) commonly assumed Internet myths
- even with use of a number of data sets, we (as a community) have quite low integrity in drawing macroscopic inferences
- **implication:**
  - the community (we) could make much better use of our collective intellectual resources
  - validate ideas against a larger variety of empirical data sets
  - before investing research and development time and energy on ideas that attempt to affect the infrastructure

now what?

- **‘seamless’ infrastructure: no such thing (right now)**
- **measurement tools/architecture**
  - well-considered
  - strategically deployed
  - collaboratively maintained
- **more operationally relevant research on resulting data**
  - feedback into tool design
- **correlation among data sources/types, simulation, visualization**
- **proactive participation**
  - top-down (app developers scope constraints)
  - bottom-up (ISP cooperation)

it is a great advantage for a system of philosophy to be substantially true.

-- george santayana

www.caida.org/outreach/presentations/

kc

ucsd/sdsc/caida

kc@caida.org

www.caida.org